

PATENT ABSTRACTS OF JAPAN

(11)Publication number : 08-314790

(43)Date of publication of application : 29.11.1996

(51)Int.Cl.

G06F 12/00

G06F 12/00

G06F 3/14

G06F 17/27

G06F 17/30

(21)Application number : 07-122274

(71)Applicant : HITACHI LTD

(22)Date of filing : 22.05.1995

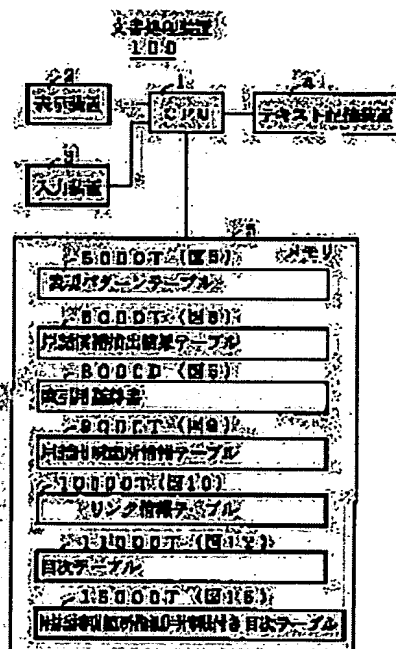
(72)Inventor : ASHIZAWA MINORU
KAJI HIROYUKI
FUJISAWA HIROMICHI

(54) HYPERTEXT STRUCTURE GENERATION SUPPORT METHOD, TEXT REFERENCE SUPPORT METHOD, INDEX TERMINOLOGY DICTIONARY GENERATION SUPPORT METHOD, AND DOCUMENT PROCESSOR

(57)Abstract:

PURPOSE: To generate a hypertext structure without a burdn of work on a user.

CONSTITUTION: A CPU 1 reads out an electronic text from a text storage device 4 to analyze it and generates a dictionary 8000D for index and a table of contents 11000T to store them in a memory 5. Items of the table of contents are displayed on a display device 2, and when the user designates one of them by an input device 3, the text main body of this designated item is displayed. When the user designates one terminology, a table of contents 15000T with information for terminology explanation part estimation is generated and is stored in the memory 5. Next, items or the table or contents are displayed where the frequency in appearance of the designated terminology, information of coincidence/noncoincidence between this terminology and an expression pattern defined in an expression pattern table 500T, and the history of reference of this terminology are added. When the user designates one of items of the table of contents, the text main body of this item is displayed, and when the user inputs the degree of satisfaction, a link information table 1000T is generated and is stored in the memory 5.



LEGAL STATUS

[Date of request for examination]

[Date of sending the examiner's decision of rejection]

[Kind of final disposal of application other than the examiner's decision of rejection or application converted registration]

[Date of final disposal for application]

[Patent number]

[Date of registration]

[Number of appeal against examiner's decision of rejection]

[Date of requesting appeal against examiner's decision of rejection]

[Date of extinction of right]

(19)日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11)特許出願公開番号

特開平8-314790

(43)公開日 平成8年(1996)11月29日

(51)Int.Cl. ⁸	識別記号	庁内整理番号	F I	技術表示箇所
G 0 6 F 12/00	5 4 7	7623-5B	G 0 6 F 12/00	5 4 7 H
	5 1 5	7623-5B		5 1 5 M
3/14	3 1 0		3/14	3 1 0 E
17/27		9288-5L	15/20	5 5 0 E
17/30		9288-5L		5 5 0 F

審査請求 未請求 請求項の数12 OL (全 15 頁) 最終頁に続く

(21)出願番号 特願平7-122274

(22)出願日 平成7年(1995)5月22日

(71)出願人 000005108

株式会社日立製作所

東京都千代田区神田駿河台四丁目6番地

(72)発明者 芦沢 実

東京都国分寺市東恋ヶ窪1丁目280番地

株式会社日立製作所中央研究所内

(72)発明者 梶 博行

東京都国分寺市東恋ヶ窪1丁目280番地

株式会社日立製作所中央研究所内

(72)発明者 藤澤 浩道

東京都国分寺市東恋ヶ窪1丁目280番地

株式会社日立製作所中央研究所内

(74)代理人 弁理士 有近 紳志郎

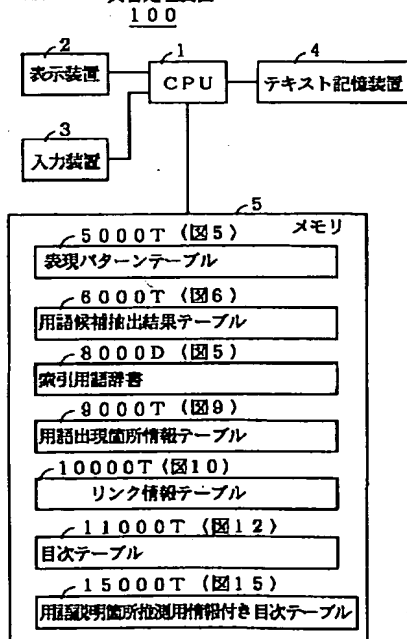
(54)【発明の名称】 ハイパーテキスト構造作成支援方法およびテキスト参照支援方法および索引用語辞書作成支援方法および文書処理装置

(57)【要約】 (修正有)

【目的】 利用者に作業負担をかけずに、ハイパーテキスト構造を作成する。

【構成】 CPU 1は、電子化テキストをテキスト記憶装置4から読み出し解析し、索引用語辞書8000Dと目次テーブル11000Tを作成し、メモリ5に記憶する。次に、目次項目を表示装置2に表示し、利用者が目次項目の一つを入力装置3により指定すると、その目次項目のテキスト本文を表示し、利用者が用語の一つを指定すると、用語説明箇所推測用情報付き目次テーブル15000Tを作成し、メモリ5に記憶する。次に、指定された用語の出現頻度と当該用語が表現パターンテーブル5000Tに定義された表現パターンと一致するか否かと当該用語が参照された履歴とを付加した目次項目を表示する。利用者が目次項目の一つを指定すると、その目次項目のテキスト本文を表示し、利用者が満足度を入力すると、リンク情報テーブル10000Tを作成し、メモリ5に記憶する。

(図1) 文書処理装置



【特許請求の範囲】

【請求項1】 テキスト本文中に使用されている用語またはその用語が使用されているテキスト本文中の箇所とその用語を説明するテキスト本文中の箇所とを結ぶリンク情報を持つハイパーテキスト構造の作成を支援するハイパーテキスト構造作成支援方法であって、

用語の指定または用語が使用されているテキスト本文中の箇所の指定を利用者から受け付け、当該用語の出現頻度による頻度重み付け、または、当該用語の出現箇所と前後の表現に基づく表現重み付け、または、当該用語が参照された履歴による履歴重み付けの少なくとも1つを目次項目ごとに求めて表示し、1以上の目次項目の指定を利用者から受け付け、前記指定された用語または指定された箇所と前記指定された目次項目に対応するテキスト本文中の箇所とを結ぶリンク情報を生成することを特徴とするハイパーテキスト構造作成支援方法。

【請求項2】 請求項1に記載のハイパーテキスト構造作成支援方法において、テキスト本文を表示するか又は予め作成した索引用語辞書に登録されている用語を一覧表示して、用語の指定または用語が使用されているテキスト本文中の箇所の指定を利用者から受け付けることを特徴とするハイパーテキスト構造作成支援方法。

【請求項3】 請求項1または請求項2に記載のハイパーテキスト構造作成支援方法において、前記指定された目次項目に対応するテキスト本文を表示し、満足度の入力を利用者から受け付け、入力された満足度に応じて前記重み付けを更新することを特徴とするハイパーテキスト構造作成支援方法。

【請求項4】 請求項2または請求項3に記載のハイパーテキスト構造作成支援方法において、テキスト本文を表示する際、予め作成した索引用語辞書に登録されている用語を強調表示することを特徴とするハイパーテキスト構造作成支援方法。

【請求項5】 請求項2から請求項4のいずれかに記載のハイパーテキスト構造作成支援方法において、テキスト本文の形態素解析を行って用語候補を抽出し、当該用語候補が目次項目のタイトルに含まれるか否かのタイトル情報、または、当該用語候補の出現頻度による頻度重み付け、または、当該用語候補の出現箇所と前後の表現に基づく表現重み付けの少なくとも1つを用語候補ごとに求めて表示し、1以上の用語候補の指定を利用者から受け付け、指定された用語候補を用語として登録して索引用語辞書を作成することを特徴とするハイパーテキスト構造作成支援方法。

【請求項6】 テキスト本文中に使用されている用語を説明する箇所のテキスト本文を参照することを支援するテキスト参照支援方法であって、

用語の指定を利用者から受け付け、当該用語の出現頻度による頻度重み付け、または、当該用語の出現箇所と前後の表現に基づく表現重み付け、または、当該用語が参

照された履歴による履歴重み付けの少なくとも1つを目次項目ごとに求めて表示し、1つの目次項目の指定を利用者から受け付け、指定された目次項目に対応するテキスト本文を表示することを特徴とするテキスト参照支援方法。

【請求項7】 請求項6に記載のテキスト参照支援方法において、テキスト本文を表示するか又は予め作成した索引用語辞書に登録されている用語を一覧表示して、用語の指定を利用者から受け付けることを特徴とするテキスト参照支援方法。

【請求項8】 請求項6または請求項7に記載のテキスト参照支援方法において、テキスト本文を表示する際、予め作成した索引用語辞書に登録されている用語を強調表示することを特徴とするテキスト参照支援方法。

【請求項9】 請求項7または請求項8に記載のテキスト参照支援方法において、テキスト本文の形態素解析を行って名詞または未知語が連続する部分を用語候補として抽出し、当該用語候補が目次項目のタイトルに含まれるか否かのタイトル情報、または、当該用語候補の出現頻度による頻度重み付け、または、当該用語候補の出現箇所と前後の表現に基づく表現重み付けの少なくとも1つを用語候補ごとに求めて表示し、1以上の用語候補の指定を利用者から受け付け、指定された用語候補を用語として登録して索引用語辞書を作成することを特徴とするテキスト参照支援方法。

【請求項10】 テキスト本文の形態素解析を行って名詞または未知語が連続する部分を用語候補として抽出し、当該用語候補が目次項目のタイトルに含まれるか否かのタイトル情報、または、当該用語候補の出現頻度による頻度重み付け、または、当該用語候補の出現箇所と前後の表現に基づく表現重み付けの少なくとも1つを用語候補ごとに求めて表示し、1以上の用語候補の指定を利用者から受け付け、指定された用語候補を用語として登録して索引用語辞書を作成することを特徴とする索引用語辞書作成支援方法。

【請求項11】 用語の指定または用語が使用されているテキスト本文中の箇所の指定を利用者から受け付ける用語箇所指定受付手段と、当該用語の出現頻度による頻度重み付け、または、当該用語の出現箇所と前後の表現に基づく表現重み付け、または、当該用語が参照された履歴による履歴重み付けの少なくとも1つを目次項目ごとに求めて表示する目次項目重み付け表示手段と、1以上の目次項目の指定を利用者から受け付ける目次項目指定受付手段と、前記指定された用語または指定された箇所と前記指定された目次項目に対応するテキスト本文中の箇所とを結ぶリンク情報を生成するリンク情報生成手段とを具備したことを特徴とする文書処理装置。

【請求項12】 用語の指定を利用者から受け付ける用語指定受付手段と、当該用語の出現頻度による頻度重み付け、または、当該用語の出現箇所と前後の表現に基づく

3

く表現重み付け、または、当該用語が参照された履歴による履歴重み付けの少なくとも1つを目次項目ごとに求めて表示する目次項目重み付け表示手段と、1つの目次項目の指定を利用者から受け付ける目次項目指定受付手段と、指定された目次項目に対応するテキスト本文を表示するテキスト本文表示手段とを具備したことを特徴とする文書処理装置。

【発明の詳細な説明】

【0001】

【産業上の利用分野】本発明は、ハイパーテキスト構造作成支援方法およびテキスト参照支援方法および索引用語辞書作成支援方法および文書処理装置に関する。さらに詳しくは、テキスト本文中に出現する用語とその説明箇所とをリンクさせたハイパーテキスト構造の作成を支援するハイパーテキスト構造作成支援方法およびテキスト本文中に出現する用語からその説明箇所を参照することを支援するテキスト参照支援方法および索引用語辞書を効率的に作成できる索引用語辞書作成方法および前記ハイパーテキスト構造作成支援方法を実施するのに好適な文書処理装置および前記テキスト参照支援方法を実施するの

【0002】

【従来の技術】索引を利用することによって、テキスト本文中に出現する用語とその説明箇所とをリンクさせたハイパーテキスト(Hyper Text)構造を作成したり、テキスト本文中に出現する用語からその説明箇所を迅速に参照することが出来る。索引の作成を支援する従来技術は、例えば「千村浩靖、佐古慎二：“マニュアル作成支援ツールの開発(2)——索引作成支援ツール——”、情報処理学会第44回(平成4年前期)全国大会、3C-2、3-259頁~3-260頁(1994)」に記載されている。これは、画面に表示されたテキスト本文上で用語を指定すると、テキスト本文をサーチして、その用語の出現位置を利用者に提示し、その提示した位置を索引とするか否かを利用者に指示させ、索引を構築していくものである。

【0003】

【発明が解決しようとする課題】上記索引の作成を支援する従来技術では、用語の出現位置の前後の文章を利用者が読んで、その位置を索引とするか否かを利用者が逐一判断しなければならない。しかし、用語の出現位置は多数あるため、すべての位置の文章を利用者が読んで索引とするか否かを判断する作業は容易でなく、利用者に作業負担がかかる問題点がある。そこで、本発明の第1の目的は、利用者に作業負担をかけることなく、テキスト本文中に出現する用語とその説明箇所とをリンクさせたハイパーテキスト構造を作成できるように支援するハイパーテキスト構造作成支援方法を提供することにある。また、本発明の第2の目的は、利用者に作業負担をかけることなく、テキスト本文中に出現する用語からそ

4

の説明箇所を参照できるように支援するテキスト参照支援方法を提供することにある。また、本発明の第3の目的は、利用者に作業負担をかけることなく、索引用語辞書を効率的に作成することができる索引用語辞書作成方法を提供することにある。また、本発明の第4の目的は、前記ハイパーテキスト構造作成支援方法を実施するのに好適な文書処理装置を提供することにある。また、本発明の第5の目的は、前記テキスト参照支援方法を実施するの

【0004】

【課題を解決するための手段】第1の観点では、本発明は、テキスト本文中に使用されている用語またはその用語が使用されているテキスト本文中の箇所とその用語を説明するテキスト本文中の箇所とを結びリンク情報を持つハイパーテキスト構造の作成を支援するハイパーテキスト構造作成支援方法であって、用語の指定または用語が使用されているテキスト本文中の箇所の指定を利用者から受け付け、当該用語の出現頻度による頻度重み付け、または、当該用語の出現箇所と前後の表現に基づく表現重み付け、または、当該用語が参照された履歴による履歴重み付けの少なくとも1つを目次項目ごとに求めて表示し、1以上の目次項目の指定を利用者から受け付け、前記指定された用語または指定された箇所と前記指定された目次項目に対応するテキスト本文中の箇所とを結びリンク情報を生成することを特徴とするハイパーテキスト構造作成支援方法を提供する。

【0005】第2の観点では、本発明は、上記構成のハイパーテキスト構造作成支援方法において、テキスト本文を表示するか又は予め作成した索引用語辞書に登録されている用語を一覧表示して、用語の指定または用語が使用されているテキスト本文中の箇所の指定を利用者から受け付けることを特徴とするハイパーテキスト構造作成支援方法を提供する。

【0006】第3の観点では、本発明は、上記構成のハイパーテキスト構造作成支援方法において、前記指定された目次項目に対応するテキスト本文を表示し、満足度の入力を利用者から受け付け、入力された満足度に応じて前記重み付けを更新することを特徴とするハイパーテキスト構造作成支援方法を提供する。

【0007】第4の観点では、本発明は、上記構成のハイパーテキスト構造作成支援方法において、テキスト本文を表示する際、予め作成した索引用語辞書に登録されている用語を強調表示することを特徴とするハイパーテキスト構造作成支援方法を提供する。

【0008】第5の観点では、本発明は、上記構成のハイパーテキスト構造作成支援方法において、テキスト本文の形態素解析を行って用語候補を抽出し、当該用語候補が目次項目のタイトルに含まれるか否かのタイトル情報、または、当該用語候補の出現頻度による頻度重み付け、または、当該用語候補の出現箇所と前後の表現に基

5

づく表現重み付けの少なくとも1つを用語候補ごとに求めて表示し、1以上の用語候補の指定を利用者から受け付け、指定された用語候補を用語として登録して索引用語辞書を作成することを特徴とするハイパーテキスト構造作成支援方法を提供する。

【0009】第6の観点では、本発明は、テキスト本文中に使用されている用語を説明する箇所のテキスト本文を参照することを支援するテキスト参照支援方法であって、用語の指定を利用者から受け付け、当該用語の出現頻度による頻度重み付け、または、当該用語の出現箇所と前後の表現に基づく表現重み付け、または、当該用語が参照された履歴による履歴重み付けの少なくとも1つを目次項目ごとに求めて表示し、1つの目次項目の指定を利用者から受け付け、指定された目次項目に対応するテキスト本文を表示することを特徴とするテキスト参照支援方法を提供する。

【0010】第7の観点では、本発明は、上記構成のテキスト参照支援方法において、テキスト本文を表示するか又は予め作成した索引用語辞書に登録されている用語を一覧表示して、用語の指定を利用者から受け付けることを特徴とするテキスト参照支援方法を提供する。

【0011】第8の観点では、本発明は、上記構成のテキスト参照支援方法において、テキスト本文を表示する際、予め作成した索引用語辞書に登録されている用語を強調表示することを特徴とするテキスト参照支援方法を提供する。

【0012】第9の観点では、本発明は、上記構成のテキスト参照支援方法において、テキスト本文の形態素解析を行って名詞または未知語が連続する部分を用語候補として抽出し、当該用語候補が目次項目のタイトルに含まれるか否かのタイトル情報、または、当該用語候補の出現頻度による頻度重み付け、または、当該用語候補の出現箇所と前後の表現に基づく表現重み付けの少なくとも1つを用語候補ごとに求めて表示し、1以上の用語候補の指定を利用者から受け付け、指定された用語候補を用語として登録して索引用語辞書を作成することを特徴とするテキスト参照支援方法を提供する。

【0013】第10の観点では、本発明は、テキスト本文の形態素解析を行って名詞または未知語が連続する部分を用語候補として抽出し、当該用語候補が目次項目のタイトルに含まれるか否かのタイトル情報、または、当該用語候補の出現頻度による頻度重み付け、または、当該用語候補の出現箇所と前後の表現に基づく表現重み付けの少なくとも1つを用語候補ごとに求めて表示し、1以上の用語候補の指定を利用者から受け付け、指定された用語候補を用語として登録して索引用語辞書を作成することを特徴とする索引用語辞書作成支援方法を提供する。

【0014】第11の観点では、本発明は、用語の指定または用語が使用されているテキスト本文中の箇所の指

6

定を利用者から受け付ける用語箇所指定受付手段と、当該用語の出現頻度による頻度重み付け、または、当該用語の出現箇所と前後の表現に基づく表現重み付け、または、当該用語が参照された履歴による履歴重み付けの少なくとも1つを目次項目ごとに求めて表示する目次項目重み付け表示手段と、1以上の目次項目の指定を利用者から受け付ける目次項目指定受付手段と、前記指定された用語または指定された箇所と前記指定された目次項目に対応するテキスト本文中の箇所とを結ぶリンク情報を生成するリンク情報生成手段とを具備したことを特徴とする文書処理装置を提供する。

【0015】第12の観点では、本発明は、用語の指定を利用者から受け付ける用語指定受付手段と、当該用語の出現頻度による頻度重み付け、または、当該用語の出現箇所と前後の表現に基づく表現重み付け、または、当該用語が参照された履歴による履歴重み付けの少なくとも1つを目次項目ごとに求めて表示する目次項目重み付け表示手段と、1つの目次項目の指定を利用者から受け付ける目次項目指定受付手段と、指定された目次項目に対応するテキスト本文を表示するテキスト本文表示手段とを具備したことを特徴とする文書処理装置を提供する。

【0016】

【作用】上記第1の観点のハイパーテキスト構造作成支援方法および上記第11の観点の文書処理装置では、利用者が指定した用語の出現頻度に基づく頻度重み付けと、当該用語の出現箇所の前後の表現に基づく表現重み付けと、当該用語が参照された履歴による履歴重み付けの少なくとも1つを目次項目ごとに求めて表示する。そして、1以上の目次項目の指定を利用者から受け付けて、リンク情報を生成する。このように、利用者は、頻度重み付け、表現重み付け、履歴重み付けの少なくとも1つを手掛かりにして、当該用語の適切な説明箇所としての蓋然性が高い目次項目を指定できる。従って、用語のすべての出現位置の文章を利用者が読む必要がなくなり、利用者に作業負担がかからなくなる。

【0017】上記第2の観点のハイパーテキスト構造作成支援方法では、テキスト本文を表示するか又は索引用語辞書に登録されている用語を一覧表示して、用語の指定または用語が使用されている箇所の指定を利用者から受け付ける。テキスト本文を表示する場合は、前後の文章を読むことが出来るので、内容を良く知らない一般読者でも容易に指定を行うことが出来る。他方、索引用語辞書に登録されている用語を一覧表示する場合は、文章を読む必要がないので、内容を良く知っている著者や編者が効率的に指定を行うことが出来る。

【0018】上記第3の観点のハイパーテキスト構造作成支援方法では、利用者が指定した目次項目に対応するテキスト本文を表示し、指定された用語または指定された箇所と前記指定された目次項目に対応するテキスト本

50

文中の箇所とを結ぶことの満足度の入力を利用者から受け付け、入力された満足度に応じて前記頻度重み付け、表現重み付け、履歴重み付けの少なくとも1つを更新する。これにより、目次項目の指定の手掛かりに対し利用者の評価をフィードバックすることが出来るので、その後の目次項目の指定の精度を向上させることが出来る。

【0019】上記第4の観点のハイパーテキスト構造作成支援方法では、テキスト本文を表示する際、索引用語辞書に登録されている用語を強調表示する。これにより、どれが用語であるかを利用者が容易に視認できるようになる。

【0020】上記第5の観点のハイパーテキスト構造作成支援方法および上記第9の観点のテキスト参照支援方法および上記第10の観点の索引用語辞書作成方法では、テキスト本文の形態素解析を行って用語候補を抽出し、当該用語候補が目次項目のタイトルに含まれるか否かのタイトル情報と、当該用語候補の出現頻度による頻度重み付けと、当該用語候補の出現箇所と前後の表現に基づく表現重み付けの少なくとも1つを用語候補ごとに求めて表示する。そして、用語候補の指定を利用者から受け付け、指定された用語候補を用語として登録して索引用語辞書を作成する。このように、利用者は、タイトル情報、頻度重み付け、表現重み付けの少なくとも1つを手掛かりにして、多数の用語候補の中から適切な用語を指定できる。従って、いちいち文章を利用者が読む必要がなくなり、利用者に作業負担がかからなくなる。

【0021】上記第6の観点のテキスト参照支援方法および上記第12の観点の文書処理装置では、利用者が指定した用語の出現頻度に基づく頻度重み付けと、当該用語の出現箇所の前後の表現に基づく表現重み付けと、当該用語が参照された履歴による履歴重み付けの少なくとも1つを目次項目ごとに求めて表示する。そして、1つの目次項目の指定を利用者から受け付けて、指定された目次項目に対応するテキスト本文を表示する。このように、利用者は、頻度重み付け、表現重み付け、履歴重み付けの少なくとも1つを手掛かりにして、当該用語の適切な説明箇所としての蓋然性が高い目次項目を指定できる。従って、すべての出現位置の文章を利用者が読む必要がなくなり、利用者に作業負担をかけずに、用語の説明箇所を参照できる。

【0022】上記第7の観点のテキスト参照支援方法では、テキスト本文を表示するか又は索引用語辞書に登録されている用語を一覧表示して、用語の指定を利用者から受け付ける。テキスト本文を表示する場合は、前後の文章を読むことが出来るので、内容を良く知らない一般読者でも容易に指定を行うことが出来る。他方、索引用語辞書に登録されている用語を一覧表示する場合は、文章を読む必要がないので、内容を良く知っている著者や編者が効率的に指定を行うことが出来る。

【0023】上記第8の観点のテキスト参照支援方法で

は、テキスト本文を表示する際、索引用語辞書に登録されている用語を強調表示する。これにより、どれが用語であるかを利用者が容易に視認できるようになる。

【0024】

【実施例】以下、図に示す実施例により本発明をさらに詳細に説明する。なお、これにより本発明が限定されるものではない。図1は、本発明の一実施例の文書処理装置100を示す構成図である。この文書処理装置100は、各部の作動を制御するCPU1と、文書処理に必要な各種の情報を画面に表示する表示装置2と、利用者からの指示を受け付ける入力装置3と、電子文書化されたテキストを記憶するテキスト記憶装置4と、プログラムやデータを記憶するメモリ5とを具備して構成されている。前記メモリ5は、表現パターンテーブル5000Tと、用語候補抽出結果テーブル6000Tと、索引用語辞書8000Dと、用語出現箇所情報テーブル9000Tと、リンク情報テーブル10000Tと、目次テーブル11000Tと、用語説明箇所推測情報付き目次テーブル15000Tとを格納する。これら各テーブルおよび辞書の内容については、後で詳しく説明する。

【0025】図2は、この文書処理装置100によるテキストの組織化処理の手順を示すフロー図である。索引用語辞書作成ステップ1001では、索引用語辞書8000Dおよび用語出現箇所情報テーブル9000Tを作成する。

【0026】図3に、上記索引用語辞書作成ステップ1001の処理手順の詳細なフロー図を示す。テキスト入力サブステップ2001では、章、節などの文書構造とそれら文書構造のタイトルがSGML (Standard Generalized Markup Language; 汎用組版指示言語) によって明示されたテキストを電子文書として入力し、テキスト記憶装置4に記憶させる。

【0027】図4に、テキストの一例を示す。タグ<DOC>4001は、文書の開始を表わす。タグ<TITLE>4002は、文書のタイトルを表わす。タグ<CHAP>4003は、章の開始を表わす。タグ<CTITLE>4004は、章のタイトルを表わす。タグ<SECT>4005は、節の開始を表わす。タグ<STITLE>は、節のタイトルを表わす。なお、“...”は、省略した図示を表わす。他の表においても同様である。

【0028】図3に戻り、用語候補抽出サブステップ2002では、入力されたテキストの形態素解析を行い、名詞または未知語が連続する部分を用語候補として抽出し、当該用語候補がタイトル(文書、章、節のタイトル)に含まれるか否かの判定結果を用語候補抽出結果テーブル6000Tに記録する。また、テキストの本文における当該用語候補の出現頻度による頻度重み付けを用語候補抽出結果テーブル6000Tに記録する。さらに、当該用語候補を含む表現パターンが表現パターンテーブル5000Tに定義された用語定義表現パターンと

一致するか否かによる表現重み付けを用語候補抽出結果テーブル6000Tに記録する。

【0029】図5に、表現パターンテーブル5000Tの一例を示す。表現パターン“\$”5001は、当該用語候補“\$”が“f”と“j”で挟まれている表現パターンである。例えば、図4のテキストにおいて、部分4007から用語候補として“概念依存図式”が抽出され、それは“f”と“j”で挟まれているので、表現パターン5001と一致する。表現パターン“~である\$”5002は、任意の文字列“~”に“である”が続き、それに用語候補“\$”が続く表現パターンである。例えば、図4のテキストにおいて、部分4009から用語候補として“日本語”が抽出され、それは“入力言語である”に続くので、表現パターン5002と一致する。

【0030】図6に、用語候補抽出結果テーブル6000Tの一例を示す。用語候補欄6002には、抽出された各用語候補を格納する。例えば、図4のテキストから抽出された“日本語”“概念依存図式”…を格納する。タイトル欄6001には、用語候補がタイトルに含まれているときに、マーク“*”を格納する。例えば、図4のテキストにおいて、用語候補“概念依存図式”は節のタイトル“2. 2 概念依存図式を用いた翻訳処理の概要”に含まれているので、エントリ6005に“*”を格納する。頻度欄6003には、用語候補の出現頻度を頻度重み付けとして格納する。例えば、用語候補“日本語”の出現頻度が24回なら“2.4”を格納する。表現欄6004には、用語候補の表現パターンが表現パターンテーブル5000T(図5参照)に定義された用語定義表現パターンと一致するなら表現重み付けとして“1”を格納し、一致しなければ“0”を格納する。例えば、用語候補“概念依存図式”は表現パターン5001と一致するので“1”を格納する。なお、用語候補と用語定義表現パターンの一致が複数回でも“1”を格納する。

【0031】図3に戻り、用語候補指定入力サブステップ2003では、索引用語とすべき重要な用語候補を利用者により指定可能とするための用語候補指定入力画面を表示装置2に表示する。

【0032】図7に、用語候補指定入力画面を例示する。この用語候補指定入力画面では、タイトル欄7001と、用語候補欄7002と、頻度欄7003と、表現欄7004と、重要チェック欄7005とを持つテーブルを表示する。また、用語候補指定入力処理を終了するためのフィールド“終了”7006を表示する。前記タイトル欄7001~表現欄7004では、用語候補抽出結果テーブル6000T(図6参照)の内容を提示する。前記重要チェック欄7005では、利用者による重要度の指定の結果を表示する。最初の表示時点では、重要チェック欄7005は、すべて“×”(重要でないこ

とを示すマーク)である。利用者が索引用語とすべき用語候補を指定すると、それに対応するエントリに“○”(特に重要であることを示すマーク)を表示する。利用者が、フィールド“終了”7006を指定すると、用語候補指定入力サブステップ2003を終了する。

【0033】なお、図7の前記重要チェック欄7005や前記フィールド“終了”7006における下線は、当該欄やフィールドの内容を利用者が指定可能であることを表している。

【0034】図3に戻り、索引用語辞書作成サブステップ2004では、索引用語辞書8000Dを作成する。図8に、上記索引用語辞書8000Dの一例を示す。索引用語辞書8000Dは、上記の用語候補指定入力サブステップ2003で重要チェック欄7005を“○”とした用語候補を集めた用語リストである。

【0035】図3に戻り、用語出現箇所情報作成サブステップ2005では、索引用語辞書8000Dの各用語の出現箇所や、各用語の前後の重み付け等を記録した用語出現箇所情報テーブル9000Tを作成する。

【0036】図9に、上記用語出現箇所情報テーブル9000Tの一例を示す。用語欄9001には、各用語を格納する。出現文書構造欄9002には、用語が出現する文書構造(文書構造はSGMLのタグとその順番で表記される)を格納する。例えば、用語“疑似的句構造”9005に対する“((DOC, 1) (CHAP, 1))”は、図4のテキストにおいて1番目のタグ<DOC>の文書の1番目のタグ<CHAP>の章に出現していることを表している。また、例えば、用語“疑似的句構造”9006に対する“((DOC, 1) (CHAP, 2) (SECT, 1))”は、図4のテキストにおいて1番目のタグ<DOC>の文書の2番目のタグ<CHAP>の章の3番目のタグ<SECT>の節に出現していることを表している。

【0037】文字位置欄9003には、前記出現文書構造欄9002で表記された文書構造の先頭から用語の出現位置までの文字数を格納する。例えば、用語“概念依存図式”9005に対する“44”は、図4のテキストにおいて1番目のタグ<DOC>の文書の1番目のタグ<CHAP>の章の44文字目に出現していることを表している。また、例えば、用語“概念依存図式”9006に対する“153”は、図4のテキストにおいて1番目のタグ<DOC>の文書の2番目のタグ<CHAP>の章の3番目のタグ<SECT>の節の153文字目に出現していることを表している。

【0038】表現欄9004には、対応する用語の表現重み付けの結果を格納する(先述の用語候補抽出サブステップ2002における表現重み付けと同様である)。例えば、用語“概念依存図式”9005は、表現パターンテーブル5000Tにおける表現パターン“\$”5001と一致するので“1”を格納する。また、例え

ば、用語“概念依存図式”9006は、表現パターンテーブル5000Tのいずれの表現パターンとも一致しないので“0”を格納する。

【0039】図2に戻り、リンク情報初期化ステップ1002では、リンク情報テーブル1000T（初期状態）を作成する。図10に、上記リンク情報テーブル1000Tの一例を示す。用語欄10001には、索引用語辞書8000D（図8参照）に登録した用語を格納する。リンク先欄10002は、初期状態では空である。後で図15を参照して説明するように、リンク先欄10002には、用語の出現文書構造（図9の出現文書構造欄9002の内容に相当）および過去に参照された累積回数が記録される。なお、リンク先欄10002には、一つの用語に対して複数の出現文書構造および累積回数を記述可能である。

【0040】図2に戻り、文書参照およびリンク情報作成ステップ1003では、文書の参照とリンクの作成を行う。

【0041】図11に、文書参照およびリンク情報作成ステップ1003の処理手順の詳細なフロー図を示す。目次テーブル作成サブステップ3001では、入力されたテキストからタイトルを抽出し、それらタイトルと本文とのリンクを記録した目次テーブル11000Tを作成する。

【0042】図12に、上記目次テーブル11000Tの一例を示す。タイトル欄11001には、文書、章、節などのタイトルを格納する。目次リンク欄11002には、タイトルに対応する文書構造を格納する。例えば、タイトル“1. はじめに”11004における目次リンク“((DOC, 1) (CHAP, 1))”11003は、図4のテキストにおいて1番目のタグ<DOC>の文書の2番目のタグ<CHAP>の章のタイトルであることを表わしている。

【0043】図11に戻り、モード設定サブステップ3002では、利用者モードを設定する。利用者モードには、文書の作成者や編集者のための編著者モードと、一般の読者のための一般読者モードの2つがある。ここでは、一般読者モードを指定した場合を想定して説明を続ける。モード判定サブステップ3003では、利用者モードを判定する。一般読者モードであれば目次表示サブステップ3004に進み、編著者モードであればモード判定サブステップ3006に進む。ここでは、一般読者モードを指定しているので、目次表示ステップ3004に進む。

【0044】目次表示サブステップ3004では、目次として目次テーブル11000T（図12参照）のタイトル欄11001の内容を表示する。図13に、目次表示画面の一例を示す。タイトル“機械翻訳システム”12002のように下線がないタイトルは、利用者による指定が出来ないことを表わす。タイトル“1. はじめ

に”12001のように下線があるタイトルは、利用者による指定が可能であることを表わす。

【0045】図11に戻り、本文表示箇所指定入力サブステップ3005では、目次表示画面中の各タイトルの中から本文内容を参照したいものを利用者が指定する。すると、その指定されたタイトルに対応する本文を注目本文とする。ここでは、タイトル“1. はじめに”12001を指定したものとして説明を続ける。

【0046】モード判定サブステップ3006では、再び利用者モードを判定する。一般読者モードであれば用語の強調付き本文表示サブステップ3007に進み、編著者モードであれば用語一覧表示サブステップ3008に進む。ここでは、一般読者モードを指定しているので、サブステップ3007に進む。

【0047】用語の強調付き本文表示サブステップ3007では、注目本文を表示すると共に、用語出現箇所情報テーブル9000T（図9参照）を参照して、注目本文に含まれている用語（図8の索引用語辞書8000Dに登録されている用語）を抽出し強調表示する。図14に、用語の強調付き本文表示画面の一例を示す。ここでは、タイトル“1. はじめに”に対応する本文が注目本文なので、目次テーブル11000T（図12参照）の目次リンク“((DOC, 1) (CHAP, 1))”11003が指す1番目の文書の1番目の章の本文を表示する。そして、用語“概念依存図式”や、用語“疑似的句構造”や、用語“格句的構造”を強調表示する。これらの用語に下線を付けているのは、利用者による指定が可能であることを表わす。また、用語の強調付き本文表示サブステップ3007の処理を終了するためのフィールド“終了”13002を表示する。

【0048】図11に戻り、用語指定入力サブステップ3009では、利用者は、内容説明を求める用語を指定する。内容説明を求める用語がなければ、利用者は、フィールド“終了”13002を指定する。ここでは、図14で用語“概念依存図式”13001を指定したものとして説明を続ける。

【0049】用語指定入力判定サブステップ3010では、用語が指定されたか否かを判定する。用語が指定されていれば、用語説明箇所推測用情報付き目次表示サブステップ3011に進む。用語が指定されていなければ（フィールド“終了”13002が指定されていれば）、文書参照およびリンク情報作成処理（ステップ1003）を終了する。ここでは、図14で用語“概念依存図式”13001が指定されているので、用語説明箇所推測用情報付き目次表示サブステップ3011に進む。

【0050】用語説明箇所推測用情報付き目次表示サブステップ3011では、指定された用語に対応した用語説明箇所推測用情報付き目次テーブル15000Tを作成する。そして、作成した用語説明箇所推測用情報付き

目次テーブル 15000T に基づいて用語説明箇所推測用情報付き目次を表示する。

【0051】図 15 に、用語 “概念依存図式” に対応した用語説明箇所推測用情報付き目次テーブル 15000T の一例を示す。タイトル欄 15001 および目次リンク欄 15002 は、図 12 の目次テーブル 11000T のタイトル欄 11001 および目次リンク欄 11002 と同じである。頻度欄 15003 には、図 9 の用語出現箇所情報テーブル 9000T を調べて、目次リンク欄 15002 の文書構造ごとに用語 “概念依存図式” の出現頻度を集計した値を格納する。表現欄 15004 には、目次リンク欄 15002 の文書構造ごとに用語 “概念依存図式” に対応する用語出現箇所情報テーブル 9000T (図 9 参照) の表現欄 9004 を調べて、“1” が少なくとも 1 箇所あれば “1” を格納し、“1” が 1 箇所もなければ “0” を格納する。履歴重み付け欄 15005 には、目次リンク欄 15002 の文書構造ごとに用語 “概念依存図式” に対応するリンク情報テーブル 10000T の参照回数の値を格納する。図 10 に示す初期状態のリンク情報テーブル 10000T では、リンク先 10002 は空であるので、“0” を格納する。

【0052】図 16 に、図 15 の用語説明箇所推測用情報付き目次テーブル 15000T に基づく用語説明箇所推測用情報付き目次の表示画面の一例を示す。この画面中の下線は、利用者による指定が可能であることを示す。また、2 重下線は、指定された用語 “概念依存図式” を示す。

【0053】図 11 に戻り、本文表示箇所指定入力サブステップ 3012 では、利用者は、用語説明箇所推測用情報付き目次 (図 16 参照) を見て、用語の説明箇所を推測し、その説明箇所に対応するタイトルを指定する。すると、その指定されたタイトルに対応する本文を注目本文とする。ここでは、利用者は、図 16 における頻度重み付け欄および表現重み付け欄 16004 を重視して、タイトル “2. 1 意味表現” 16002 を指定したものであるとして説明を続ける。このように、利用者は、頻度重み付け及び表現重み付けを手掛かりに出来るので、用語の説明箇所を容易かつ的確に推定できる。

【0054】指定用語の強調付き本文表示サブステップ 3013 では、注目本文を表示すると共に、指定された用語を強調表示する。また、利用者の満足度 (表示内容が用語の説明箇所として適切か否かの評価) を入力するフィールドを表示する。図 17 に、指定用語の強調付き本文表示画面の一例を示す。ここでは、タイトル “2. 1 意味表現” 16002 を指定しているので、用語説明箇所推測用情報付き目次テーブル 15000T (図 15 参照) の目次リンク “((DOC, 1) (CHAP, 2) (SECT, 1))” 15006 が指す 1 番目の文書の 2 番目の章の 1 番目の節の本文を表示する。そして、指定されている用語 “概念依存図式” を強調表

示する (2 重下線によって強調表示している)。また、利用者の満足度を入力するフィールド 17002 を表示する。満足度を表わす指標値 “+1. 0” ~ “-1. 0” に下線を付けているのは、利用者による指定が可能であることを表わす。

【0055】図 11 に戻り、満足度指定入力サブステップ 3014 では、利用者は満足度を表わす指標値のいずれかを指定する。ここでは、利用者が “+1. 0” を指定したものであるとして説明を続ける。リンク情報記録サブステップ 3015 では、リンク情報テーブル 10000T の指定されている用語に対応するリンク先欄 10002 の内容を更新する。すなわち、そのリンク先欄 10002 に、表示している本文の文書構造および指定された満足度を表わす指標値を格納する。表示している本文の文書構造が既に格納されていれば、既に格納されている満足度を表わす指標値に、指定された満足度を表わす指標値を加算する。この後、上記のモード判定サブステップ 3006 に移行する。図 18 に、リンク先欄 10002 の内容を更新されたリンク情報テーブル 10000T を示す。図 17 の画面で利用者が “+1. 0” を指定したため、用語 “概念依存図式” に対応するリンク先欄 10002 の内容が “(1, (DOC, 1) (CHAP, 2) (SECT, 1)) 18001 に更新されている。

【0056】図 11 に戻り、リンク情報記録サブステップ 3015 から移行したモード判定サブステップ 3006 では、利用者モードが “一般読者モード” であるから、用語の強調付き本文表示サブステップ 3007 に進む。用語の強調付き本文表示サブステップ 3007 では、注目本文を表示すると共に、用語出現箇所情報テーブル 9000T (図 9 参照) を参照して、注目本文に含まれている用語 (図 8 の索引用語辞書 8000D に登録されている用語) を抽出し強調表示する。図 19 に、用語の強調付き本文表示画面の一例を示す。ここでは、タイトル “2. 1 意味表現” に対応する本文が注目本文なので、目次テーブル 11000T (図 12 参照) の目次リンク “((DOC, 1) (CHAP, 2) (SECT, 1))” が指す 1 番目の文書の 2 番目の章の 1 番目の節の本文を表示する。そして、用語 “意味表現” や “概念依存図式” を強調表示する。これらの用語に下線を付けているのは、利用者による指定が可能であることを表わす。また、文書参照およびリンク情報作成処理 (ステップ 1003) を終了するためのフィールド “終了” 19001 を表示する。

【0057】図 11 に戻り、用語指定入力サブステップ 3009 では、利用者は、内容説明を求める用語を指定する。内容説明を求める用語がなければ、利用者は、フィールド “終了” 19001 を指定する。ここでは、図 19 で用語 “概念依存図式” 19002 を指定したものであるとして説明を続ける。

【0058】用語指定入力判定サブステップ 3010 で

は、用語が指定されたか否かを判定する。用語が指定されていれば、用語説明箇所推測用情報付き目次表示サブステップ3011に進む。用語が指定されていなければ(フィールド“終了”19002が指定されていれば)、文書参照およびリンク情報作成処理(ステップ1003)を終了する。ここでは、図19で用語“概念依存図式”19002が指定されているので、用語説明箇所推測用情報付き目次表示サブステップ3011に進む。

【0059】用語説明箇所推測用情報付き目次表示サブステップ3011では、指定された用語に対応した用語説明箇所推測用情報付き目次テーブル15000Tを作成する。そして、作成した用語説明箇所推測用情報付き目次テーブル15000Tに基づいて用語説明箇所推測用情報付き目次を表示する。

【0060】図20に、用語“概念依存図式”に対応した用語説明箇所推測用情報付き目次テーブル15000Tの一例を示す。タイトル欄15001、目次リンク欄15002、頻度欄15003、表現欄15004は、図15と同じである。履歴重み付け欄15005は、図18に示す更新されたリンク情報テーブル10000Tでは、リンク先18001に参照回数“1”が格納されているので、“1”を格納する。図21に、図20の用語説明箇所推測用情報付き目次テーブル15000Tに基づく用語説明箇所推測用情報付き目次の表示画面の一例を示す。

【0061】図11に戻り、本文表示箇所指定入力サブステップ3012では、利用者は、用語説明箇所推測用情報付き目次(図21参照)を見て、用語の説明箇所を推測し、その説明箇所に対応するタイトルを指定する。すると、その指定されたタイトルに対応する本文を注目本文とする。ここでは、利用者は、図21における頻度重み付け欄および表現重み付け欄および履歴重み付け欄16005を重視して、タイトル“2. 1 意味表現”16002を指定したものと説明を続ける。このように、利用者は、頻度重み付けおよび表現重み付け及び履歴重み付けを手掛かりに出来るので、用語の説明箇所を容易かつ的確に推定できる。

【0062】指定用語の強調付き本文表示サブステップ3013では、注目本文を表示すると共に、指定された用語を強調表示する。また、利用者の満足度(表示内容が用語の説明箇所として適切か否かの評価)を入力するフィールドを表示する。図22に、指定用語の強調付き本文表示画面の一例を示す。ここでは、タイトル“2. 1 意味表現”16002を指定しているので、用語説明箇所推測用情報付き目次テーブル15000T(図20参照)の目次リンク“((DOC, 1) (CHAP, 2) (SECT, 1))”15007が指す1番目の文書の2番目の章の1番目の節の本文を表示する。そして、指定されている用語“概念依存図式”を強調表

示する。また、利用者の満足度を入力するフィールド17002を表示する。

【0063】図11に戻り、満足度指定入力サブステップ3014では、利用者は満足度を表わす指標値のいずれかを指定する。ここでは、利用者が“+1. 0”を指定したものと説明を続ける。リンク情報記録サブステップ3015では、リンク情報テーブル10000Tの指定されている用語に対応するリンク先欄10002の内容を更新する。すなわち、そのリンク先欄10002に、表示している本文の文書構造および指定された満足度を表わす指標値を格納する。表示している本文の文書構造が既に格納されていれば、既に格納されている満足度を表わす指標値に、指定された満足度を表わす指標値を加算する。この後、上記のモード判定サブステップ3006に移行し、用語指定入力サブステップ3009で利用者がフィールド“終了”を指定するまで、上記動作を繰り返す。図23に、リンク先欄10002の内容を更新されたリンク情報テーブル10000Tを示す。図22の画面で利用者が“+1. 0”を指定したため、用語“概念依存図式”に対応するリンク先欄10002の内容が“((DOC, 1) (CHAP, 2) (SECT, 1))”23001に更新されている。

【0064】次に、図11のモード設定サブステップ3002において、利用者が編著者モードを指定した場合について説明する。モード判定サブステップ3003では、編著者モードが指定されていると、サブステップ3004およびサブステップ3005をスキップして、モード判定サブステップ3006に進む。モード判定サブステップ3006では、編著者モードが指定されていると、用語の一覧表示サブステップ3008に進む。用語の一覧表示サブステップ3008では、索引用語辞書8000D(図8参照)の内容を読み出して、索引用語辞書登録語の一覧を表示する。

【0065】図24に、用語の一覧表示画面の一例を示す。用語20001は、索引用語辞書8000D(図8参照)に登録された用語である。用語に下線を付けているのは、利用者による指定が可能であることを表わす。また、文書参照およびリンク情報作成処理(ステップ1003)を終了するためのフィールド“終了”20002を表示する。一般に編著者は本文の内容を熟知しているので、一般読者モードのように本文を表示せず、用語を直接表示する方が、効率を向上することが出来る。

【0066】図11に戻り、用語指定入力サブステップ3009では、図24の用語の一覧表示画面の中から用語を指定する。例えば、用語“概念依存図式”20003を指定する。用語指定入力判定サブステップ3010～リンク情報記録サブステップ3015の処理は、“一般読者モード”の場合と同じである。例えば、図25に示すようなリンク情報テーブル10000Tがこれまでに作成されておれば、図26に示すような用語説明箇所

推測用情報付き目次テーブル15000Tが作成され、図27に示すような用語説明箇所推測用情報付き目次が表示されることになる。

【0067】以上の文書処理装置100によれば、次の効果が得られる。

(1) 図7に示すように、タイトル欄7001と頻度重み付け欄7003と表現重み付け欄7004とを手掛かりにして多数の用語候補7002の中から適切な用語を指定できるから、利用者に作業負担をかけずに、索引用語辞書8000Dを作成できる。

(2) 図14に示すように、テキスト本文の表示から用語を指定できるため、内容を良く知らない一般読者でも容易に指定を行うことが出来る。また、索引用語辞書8000Dに登録されている用語を強調表示するから、どれが用語であるかを利用者が容易に視認できる。

(3) 図24に示すように、用語の一覧表示から用語を指定できるため、内容を良く知っている著者や編者は効率よく指定を行うことが出来る。

(4) 図21に示すように、頻度重み付け欄16003と表現重み付け欄16004と履歴重み付け欄16005とを手掛かりにして目次項目を指定できるから、利用者に作業負担をかけずに、用語の説明箇所を参照できる(図22参照)と共に、リンク情報(テーブル)10000Tを生成できる。

(5) 図17に示すように、フィールド17002により満足度を入力し、これを履歴重み付け欄16005にフィードバックできるので、その後の目次項目の指定の精度を向上させることが出来る。

【0068】他の実施例として、編著者モードでの満足度の指標値には一般読者モードでの満足度の指標値よりも大きな重みを付けるようにしてもよい。これにより、編著者の評価を一般読者の評価よりも重視することが出来る。また、用語説明箇所推測用情報付き目次表示サブステップ3011(図11参照)において、用語説明箇所推測用情報付き目次テーブル15000T(図20参照)を作成したとき、履歴重み付け欄15005の中で適当な値を超えるものが1つしかない場合には、用語説明箇所推測用情報付き目次(図21参照)を表示せず、その1つに対応する目次項目の指示があったものと看做して本文表示箇所指定入力サブステップ3012をスキップし、指定用語の強調付き本文表示サブステップ3013へ進むようにしてもよい。これによって、一般的なハイパーテキストと同様に、テキスト本文中の用語を指定すれば、その説明のテキスト本文を直ちに参照できる。また、リンク情報テーブル10000T(図10、図18、図23、図25参照)のリンク先欄10002に、文書構造(文書、章、節など)を記録するのに加えて、利用者が文書構造の先頭からスクロールして表示した段落、文字位置を記録してもよい。

【0069】

【発明の効果】本発明のハイパーテキスト構造作成支援方法および文書処理装置によれば、利用者に作業負担をかけることなく、テキスト本文中に出現する用語とその説明箇所とをリンクさせたハイパーテキスト構造を作成できるようになる。また、本発明のテキスト参照支援方法および文書処理装置によれば、利用者に作業負担をかけることなく、テキスト本文中に出現する用語からその説明箇所を参照できるようになる。また、本発明の索引用語辞書作成方法によれば、利用者に作業負担をかけることなく、索引用語辞書を効率的に作成できるようになる。

【図面の簡単な説明】

【図1】本発明の一実施例の文書処理装置を示す構成図である。

【図2】図1の文書処理装置によるテキストの組織化処理の手順を示すフロー図である。

【図3】索引用語辞書作成ステップの詳細を示すフロー図である。

【図4】入力テキストの例示図である。

【図5】表現パターンテーブルの例示図である。

【図6】用語候補抽出結果テーブルの例示図である。

【図7】用語候補指定入力画面の例示図である。

【図8】索引用語辞書の例示図である。

【図9】用語出現箇所情報テーブルの例示図である。

【図10】リンク情報テーブル(初期状態)の例示図である。

【図11】文書参照およびリンク情報作成処理の詳細を示すフロー図である。

【図12】目次テーブルの例示図である。

【図13】目次表示の画面の例示図である。

【図14】本文表示の画面の例示図である。

【図15】用語説明箇所推測用情報付き目次テーブルの例示図である。

【図16】用語説明箇所推測用情報付き目次表示の画面の例示図である。

【図17】指定用語の強調付き本文表示の画面の例示図である。

【図18】満足度入力後のリンク情報テーブルの例示図である。

【図19】本文表示の画面の別の例示図である。

【図20】用語説明箇所推測用情報付き目次テーブルの別の例示図である。

【図21】用語説明箇所推測用情報付き目次表示の画面の別の例示図である。

【図22】指定用語の強調付き本文表示の画面の例示図である。

【図23】満足度入力後のリンク情報テーブルの別の例示図である。

【図24】用語の一覧表示画面の例示図である。

【図25】リンク情報テーブルのさらに別の例示図であ

る。

【図26】用語説明箇所推測用情報付き目次テーブルのさらに別の例示図である。

【図27】用語説明箇所推測用情報付き目次の表示画面のさらに別の例示図である。

【符号の説明】

100 文書処理装置
1 CPU
2 表示装置
3 入力装置

4

テキスト記憶装置

5

メモリ

5000T

表現パターンテーブル

6000T

用語候補抽出結果テーブル

8000D

索引用語辞書

9000T

用語出現箇所情報テーブル

10000T

リンク情報テーブル

11000T

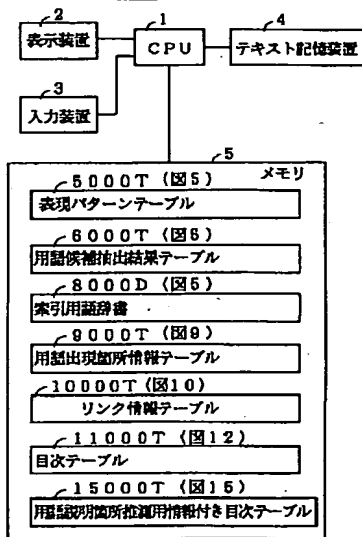
目次テーブル

15000T

用語説明箇所推測用情報付き目次テーブル

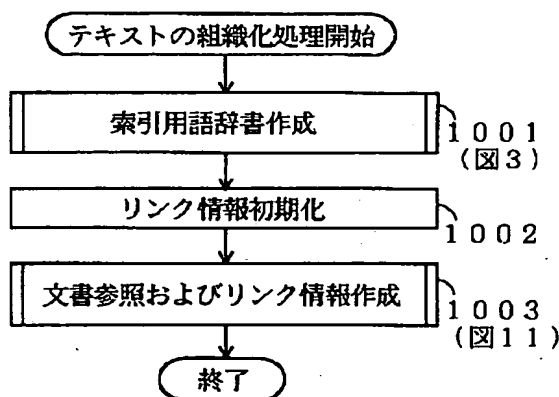
10 ル

【図1】

(図1) 文書処理装置
100

【図2】

(図2)



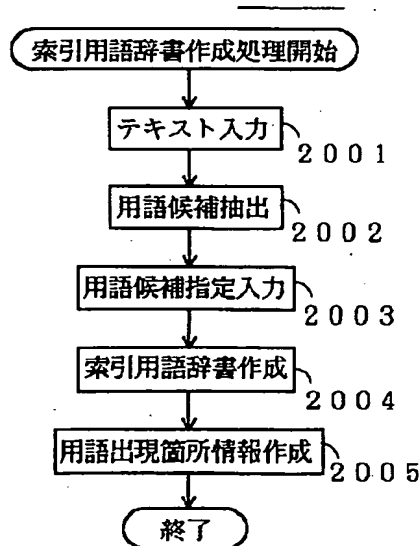
【図8】

(図8) 索引用語辞書
8000D

用語
意味処理
意味表現
英日機械翻訳システム
概念依存図式
格的句構造
疑似的句構造
日英機械翻訳システム
...

【図3】

(図3) 1001



【図5】

(図5) 表現パターンテーブル
5000T

用語定義表現パターン
5001 「る」
5002 ると名付けた
5003 るは、〜である。
5004 〜であるる
5005 るとは、
...

【図10】

(図10) リンク情報テーブル (初期状態)
10000T

用語	リンク先
意味処理	
意味表現	
英日機械翻訳システム	
概念依存図式	
格的句構造	
疑似的句構造	
日英機械翻訳システム	
...	

【図13】

機械翻訳システム	
1. はじめに	
2. 日英機械翻訳システム	
2.1 意味表現	
2.2 概念依存図式を用いた意味処理の概要	
2.3 意味処理による解析精度の向上	
3. 英日機械翻訳システム	
...	

【図4】

(図4)

4001	<DOC>	4007
4002	<TITLE> 機械翻訳システム </TITLE>	
4003	<CHAP> <CTITLE> 1. はじめに </CTITLE>	
4004	我々は、日英と英日の二つの機械翻訳システムを...	4008
	...我々の日英機械翻訳システムは、「概念依存図式」と名付けた意味表現を介した機械翻訳システムである。これは、入力言語である日本語が...	
	英日翻訳システムは、英語に直した「疑似的句構造」から日本語に適した「格的句構造」へのトランスファを行う機械翻訳システムである。これは、...	
	</CHAP>	
4005	<CHAP> <CTITLE> 2. 日英機械翻訳システム </CTITLE>	
4006	<SECT> <STITLE> 2. 1 日英翻訳における意味表現 </STITLE>	
	日本語と英語は...	
	...右向グラフを意味表現として用いることにし、これを概念依存図式と名付けた。概念依存図式では、ノードが...	4008
	</SECT>	
	<SECT> <STITLE> 2. 2 概念依存図式を用いた翻訳処理の概要 </STITLE>	
	...	
	</SECT>	
	<SECT> <STITLE> 2. 3 意味処理による解析精度の向上 </STITLE>	
	</SECT>	
	</CHAP>	
	<CHAP> <CTITLE> 3. 英日機械翻訳システム </CTITLE>	
	<SECT> <STITLE> 3. 1 英日翻訳における内部表現 </STITLE>	
	...	
	</SECT>	
	</CHAP>	
	</DOC>	

【図6】

(図6) 用語候補抽出結果テーブル
6000T

6001	6002	6003	6004
タイトル	用語候補	頻度	表現
日本語		24	1
* 概念依存図式		14	1
* 英日		10	0
* 機械翻訳システム		10	0
英語		9	0
疑似的句構造		9	1
格的句構造		8	1
* 意味表現		7	0
* 日英機械翻訳システム		6	0
トランスファ		5	0
* 英日機械翻訳システム		5	0
日英翻訳		5	0
* 意味処理		4	0
* 解析精度		4	0
...

【図7】

(図7)

7001	7002	7003	7004	7005	終了
タイトル	用語候補	頻度	表現	重要チェック	
日本語		24	1	×	
* 概念依存図式		14	1	○	
* 英日		10	0	×	
* 機械翻訳システム		10	0	×	
英語		9	0	×	
疑似的句構造		9	1	○	
格的句構造		8	1	○	
* 意味表現		7	0	○	
* 日英機械翻訳システム		6	0	○	
トランスファ		5	0	×	
* 英日機械翻訳システム		5	0	○	
日英翻訳		5	0	×	
* 意味処理		4	0	○	
* 解析精度		4	0	×	
...	

【図9】

(図9) 用語出現箇所情報テーブル
9000T

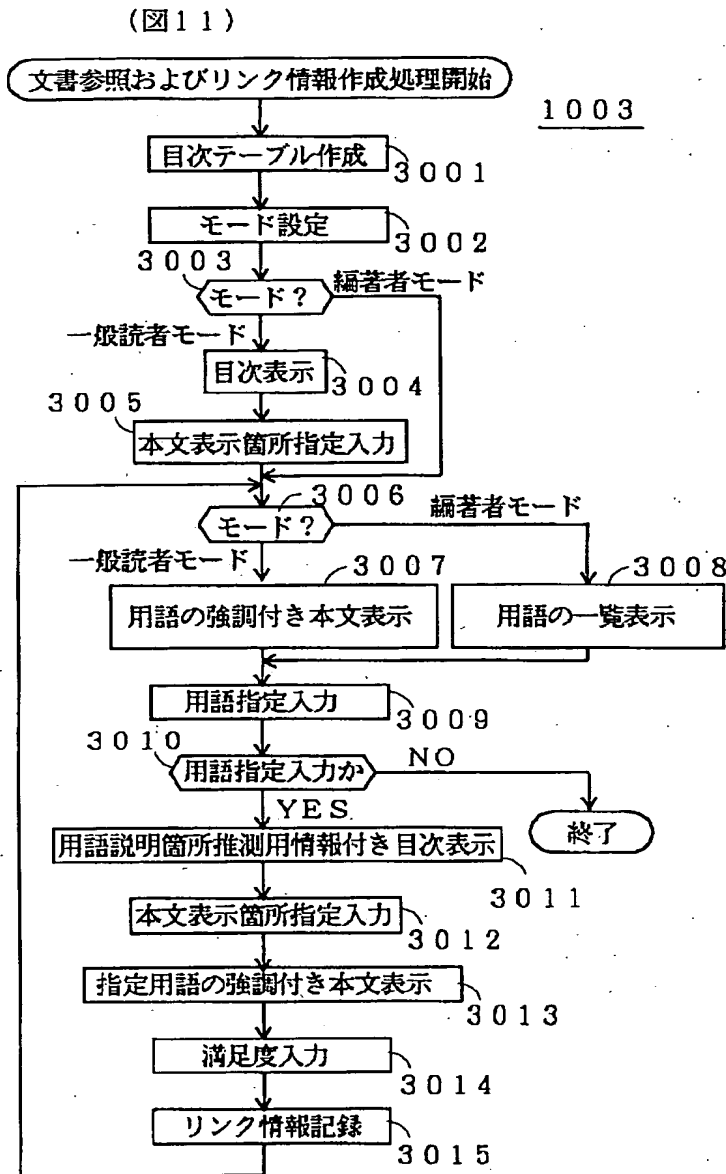
9001	9002	9003	9004
用語	出現文書構造	文字位置	表現
疑似的句構造	((DOC, 1) (CHAP, 1))	1 3 2	1
疑似的句構造	((DOC, 1) (CHAP, 3) (SECT, 1))	1 4 3	0
...	0
概念依存図式	((DOC, 1) (CHAP, 1))	4 4	1
概念依存図式	((DOC, 1) (CHAP, 2) (SECT, 1))	1 4 1	1
概念依存図式	((DOC, 1) (CHAP, 2) (SECT, 1))	1 5 3	0
...
意味表現	((DOC, 1) (CHAP, 1))	5 6	0
意味表現	((DOC, 1) (CHAP, 2) (SECT, 1))	1 3	0
意味表現	((DOC, 1) (CHAP, 2) (SECT, 1))	1 2 3	0
...

【図12】

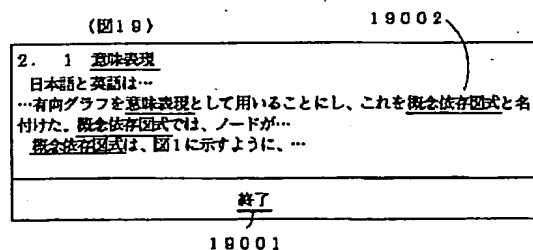
(図12) 目次テーブル 11000T

11001	11002	11003
タイトル	目次リンク	
機械翻訳システム	((DOC, 1))	
1. はじめに	((DOC, 1) (CHAP, 1))	
2. 日英機械翻訳システム	((DOC, 1) (CHAP, 2))	
2. 1 意味表現	((DOC, 1) (CHAP, 2) (SECT, 1))	
2. 2 概念依存図式を用いた翻訳処理の概要	((DOC, 1) (CHAP, 2) (SECT, 2))	
2. 3 意味処理による解析精度の向上	((DOC, 1) (CHAP, 2) (SECT, 3))	
3. 英日機械翻訳システム	((DOC, 1) (CHAP, 3))	
...	...	
...	...	

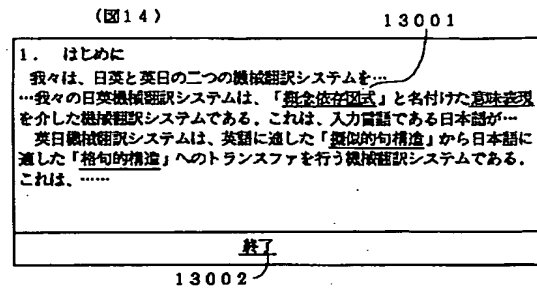
【図11】



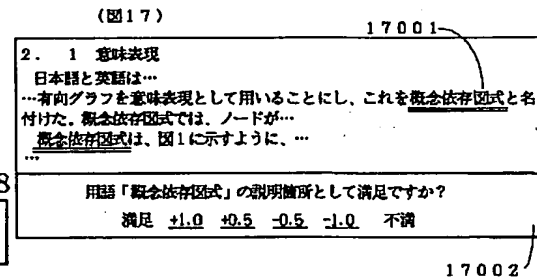
【図19】



【図14】



【図17】

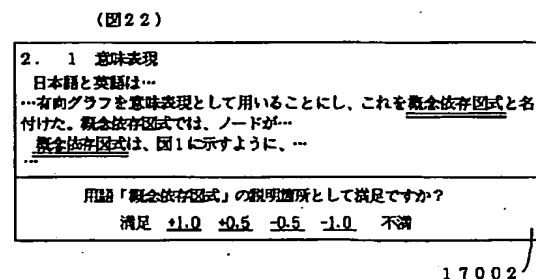


【図18】

(図18) リンク情報テーブル 10000T

用語	リンク先
意味処理	
意味表現	
英日機械翻訳システム	
概念依存図式	(1, (DOC, 1) (CHAP, 2) (SECT, 1))
格句的構造	
疑似的句構造	
日英機械翻訳システム	
…	

【図22】



【図15】

(図15) 用語説明箇所推測用情報付き目次テーブル
15000T

15001	用語	概念依存図式	15002	15003	15004	15005
	タイトル		目次リンク	頻度	表現	履歴
	機械翻訳システム		((DOC, 1))	0	0	0
1.	はじめに		((DOC, 1) (CHAP, 1))	1	1	0
2.	日英機械翻訳システム		((DOC, 1) (CHAP, 2))	0	0	0
2.1	意味表現		((DOC, 1) (CHAP, 2) (SECT, 1))	5	1	0
2.2	概念依存図式を用いた翻訳処理の概要		((DOC, 1) (CHAP, 2) (SECT, 2))	5	0	0
2.3	意味処理による解析精度の向上		((DOC, 1) (CHAP, 2) (SECT, 3))	0	0	0
3.	英日機械翻訳システム		((DOC, 1) (CHAP, 3))	0	0	0
...
...

【図16】

(図16)

用語：概念依存図式		頻度	表現	履歴
機械翻訳システム				
1.	はじめに	1	1	0
2.	日英機械翻訳システム			
2.1	意味表現	5	1	0
2.2	<u>概念依存図式を用いた翻訳処理の概要</u>	5		
2.3	<u>意味処理による解析精度の向上</u>			
3.	英日機械翻訳システム			
...	...			
...	...			

【図20】

(図20) 用語説明箇所推測用情報付き目次テーブル
15000T

15001	用語	概念依存図式	15002	15003	15004	15005
	タイトル		目次リンク	頻度	表現	履歴
	機械翻訳システム		((DOC, 1))	0	0	0
1.	はじめに		((DOC, 1) (CHAP, 1))	1	1	0
2.	日英機械翻訳システム		((DOC, 1) (CHAP, 2))	0	0	0
2.1	意味表現		((DOC, 1) (CHAP, 2) (SECT, 1))	5	1	1
2.2	概念依存図式を用いた翻訳処理の概要		((DOC, 1) (CHAP, 2) (SECT, 2))	5	0	0
2.3	意味処理による解析精度の向上		((DOC, 1) (CHAP, 2) (SECT, 3))	0	0	0
3.	英日機械翻訳システム		((DOC, 1) (CHAP, 3))	0	0	0
...
...

【図21】

(図21)

(図21)

16003 16005
16004

用語	概念依存図式	頻度	表現	履歴
機械翻訳システム				
1. はじめに		1	1	0
2. 日英機械翻訳システム				
2.1 意味表現		5	1	1
2.2 概念依存図式を用いた翻訳処理の概要		5		
2.3 意味処理による解析精度の向上				
3. 英日機械翻訳システム				
...				
...				

16002

【図23】

(図23) リンク情報テーブル 10000T

用語	リンク先
意味処理	
意味表現	
英日機械翻訳システム	
概念依存図式	(2, (DOC, 1) (CHAP, 2) (SECT, 1))
格的句構造	
疑似的句構造	
日英機械翻訳システム	
...	

【図24】

(図24)

20001	20002
用語	終了
意味処理	
意味表現	
英日機械翻訳システム	
概念依存図式	
格的句構造	
疑似的句構造	
日英機械翻訳システム	
...	

【図27】

(図27)

用語	概念依存図式	頻度	表現	履歴
機械翻訳システム				
1.	はじめに	1	1	2
2.	日英機械翻訳システム			
2.1	意味表現	5	1	10
2.2	概念依存図式を用いた翻訳処理の概要	5		
2.3	意味処理による解析精度の向上			
3.	英日機械翻訳システム			
...	...			
...	...			

【図25】

(図25) リンク情報テーブル 10000T

用語	リンク先
意味処理	...
意味表現	...
英日機械翻訳システム	...
概念依存図式	(2, (DOC, 1) (CHAP, 1)), (10, (DOC, 1) (CHAP, 2) (SECT, 1))
格的句構造	...
疑似的句構造	...
日英機械翻訳システム	...
...	...

【図26】

(図26) 用語説明箇所検索用情報付き目次テーブル 15000T

用語	概念依存図式	目次リンク	頻度	出現	位置
機械翻訳システム		((DOC, 1))	0	0	0
1. はじめに		((DOC, 1) (CHAP, 1))	1	1	2
2. 日英機械翻訳システム		((DOC, 1) (CHAP, 2))	0	0	0
2.1 意味表現		((DOC, 1) (CHAP, 2) (SECT, 1))	5	1	10
2.2 概念依存図式を用いた翻訳処理の概要		((DOC, 1) (CHAP, 2) (SECT, 2))	5	0	0
2.3 意味処理による解析精度の向上		((DOC, 1) (CHAP, 2) (SECT, 3))	0	0	0
3. 英日機械翻訳システム		((DOC, 1) (CHAP, 3))	0	0	0
...
...

フロントページの続き

(51) Int. Cl.⁶

識別記号

庁内整理番号

FI

技術表示箇所

9194-5L

G 06 F 15/419

320